AD-A042 691 WISCONSIN UNIV MADISON MATHEMATICS RESEARCH CENTER F/G 9/2

A DEFENSE OF THE KARST ALGORITHM FOR FINDING THE LINE OF BEST F--ETC(U)

MAR 77 D R SCHUETTE DAAG29-75-C-0024

UNCLASSIFIED MRC-TSR-1735 NL

| OF |

AD A042691

END

DATE FILMED

8 -77

DDC

ADA042691

MRC Technical Summary Report #1735

A DEFENSE OF THE KARST ALGORITHM FOR FINDING THE LINE OF BEST FIT UNDER THE L_1 NORM

Donald R. Schuette

12

B.S.

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

March 1977

(Received February 16, 1977)

See 1473

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

A DEFENSE OF THE KARST ALGORITHM FOR FINDING THE LINE OF BEST FIT UNDER THE L_1 NORM

Donald R. Schuette

Technical Summary Report # 1735
March 1977

ABSTRACT

The methods proposed by Karst and Sharpe of fitting a line to a given set of points in the plane so as to minimize the sum of the absolute values of the deviations are examined by means of their linear programming formulations. It is seen that the Karst algorithm, which was criticized by Sharpe, already contains the type of improvement proposed by Barrodale and Roberts for ℓ_1 linear approximations. For the single purpose of finding the optimal line the Karst procedure appears to be very efficient. If, as Sharpe suggests may sometimes be the case, the investigator is interested in the sensitivity of the minimum sum to changes in the slope parameter, then Sharpe's algorithm is preferred. The Karst algorithm is improved by incorporating into it the simplex stopping rule. The problem is generalized to permit arbitrary weightings of the deviations.

AMS (MOS) Subject Classification: 90C05

Key Words: Linear curve fitting L_1 norm, Linear programming

Work Unit Number 5 (Mathematical Programming and Operations Research)

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

A DEFENSE OF THE KARST ALGORITHM FOR FINDING THE LINE OF BEST FIT UNDER THE L_1 NORM

Donald R. Schuette

1. Introduction

Curve fitting so as to minimize the sum of the absolute deviations, i.e. under the ℓ_1-norm, has received the attention of a number of writers in recent years (e.g. [1], [3], [7], [8]). In particular, Karst [4] and Sharpe [9] have proposed algorithms for fitting a line to a given set of points in the plane under that criterion. Rao and Srinivasan [5] have interpreted Sharpe's method as the solution to the parametric dual to the linear programming formulation of the problem and have offered an alternate procedure parameterizing on the intercept rather than on the slope as in the case of Sharpe.

In this paper the linear programming formulations of Karst's and Sharpe's algorithms will be examined and compared. It will be seen that one iteration of the Karst algorithm is equivalent to several simplex iterations in the same manner as in the "*improved algorithm*" of Barrodale and Roberts. The Karst algorithm will be improved by incorporating into it the simplex stopping rule. The problem will be generalized to allow for assignment of arbitrary weights to the deviations at each point.

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

2. Geometric Considerations

Given a set of points (X_i, Y_i) and non-negative weights w_i for $i = 1, 2, \ldots, N$, the problem is to find values of the parameters a and b so as to minimize

$$S(a, b) = \sum_{i=1}^{N} w_i \left| Y_i - (a + bX_i) \right| .$$

In addition to the (X, Y) plane in which the given points lie, it is convenient to consider as did Karst and Sharpe the (a, b) plane or parameter space for the problem. The surface $S(a, b)$ lying above the (a, b) plane is piecewise linear and convex as noted by Sharpe and Karst (see also Rice [6]). Each point in the (a, b) plane corresponds to a line in the (X, Y) plane. The line $a = r - sb$ in the (a, b) plane corresponds to the family of lines in the (X, Y) plane through the point with coordinates (s, r), and the line $b = b'$ in the (a, b) plane corresponds to the family of lines in the (X, Y) plane with constant slope. Hence, each of the given points (X_i, Y_i) determines a line $a = Y_i - X_i b$ in the (a, b) plane. Such lines will be referred to as basic lines (see Figure 1).

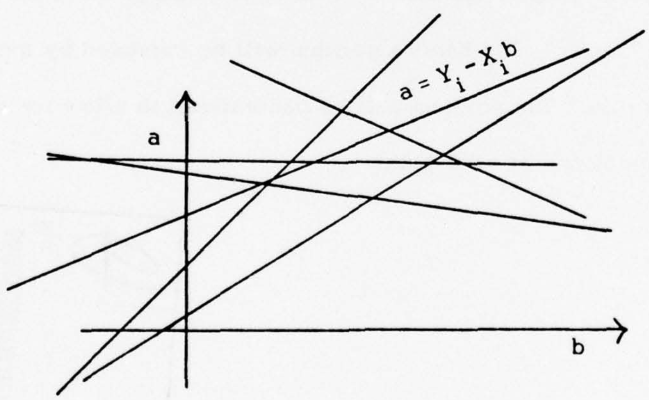

Figure 1

One of the well known results of ℓ_1 linear approximation theory (Barrodale and Roberts [3], Rice [6]) is that there is an *optimal solution which interpolates* at least two of the data points. That means that there is an optimal point in the (a, b) plane which lies at the intersection of two basic lines. When there is only one optimal point it must lie at the intersection of two basic lines. Hence, the justification for solution search procedures which restrict the search to the points of intersection of basic lines.

3. Linear Programming Formulations

It is well known that the problem of choosing a and b so as to minimize

$$S(a,b) = \sum_{i=1}^{N} w_i |Y_i - (a + bX_i)| \tag{1}$$

can be formulated as a linear programming problem (Wagner [12], Barrodale and Roberts [3], Armstrong and Frome [1]). Armstrong and Frome, for example, in the case when $w_i = 1$ for all i, proceed by setting

$$P_i - N_i = Y_i - (a + bX_i)$$

with $P_i \geq 0$ and $N_i \geq 0$ for $i = 1, 2, \ldots, N$. Thus, the vertical deviation of each point from the line with parameters a and b is divided into positive and negative components. The linear programming problem then becomes

$$\text{minimize} \quad \sum_{i=1}^{N} (P_i + N_i) \tag{2}$$

subject to

$$P_i - N_i + a + bX_i = Y_i \tag{3}$$

$$P_i \geq 0,\ N_i \geq 0 \quad \text{for} \quad i = 1, 2, \ldots, N\,,$$

with a and b unrestricted in sign.

The dual problem (see for example Wagner [12]) is

$$\text{maximize} \quad \sum_{i=1}^{N} Y_i v_i \tag{4}$$

subject to

$$\sum_{i=1}^{N} v_i = 0 \tag{5}$$

$$\sum_{i=1}^{N} X_i v_i = 0 \tag{6}$$

$$-1 \leq v_i \leq 1, \qquad i = 1, 2, \ldots, N\,. \tag{7}$$

The dual is a bounded variable linear programming problem for which special solution methods are available (see Taha [11]). If one prefers, the substitution $t_i = v_i + 1$ can be made so that the problem becomes

$$\text{maximize} \quad \sum_{i=1}^{N} Y_i t_i - \sum_{i=1}^{N} Y_i \tag{8}$$

subject to

$$\sum_{i=1}^{N} t_i = N \tag{9}$$

$$\sum_{i=1}^{N} X_i t_i = \sum_{i=1}^{N} X_i \tag{10}$$

$$0 \le t_i \le 2, \qquad i = 1,2,\ldots,n\ . \tag{11}$$

The variables are then non-negative and upper bounded.

In the more general case where the deviations are weighted by the factors w_i the primal problem is

$$\text{minimize} \quad \sum_{i=1}^{N} w_i(P_i + N_i) \tag{12}$$

subject to

$$P_i - N_i + a + bX_i = Y_i \tag{13}$$

$$P_i \ge 0,\ N_i \ge 0, \qquad i = 1,2,\ldots,N \tag{14}$$

with a and b unrestricted in sign. The dual then becomes

$$\text{maximize} \quad \sum_{i=1}^{N} Y_i v_i \tag{15}$$

subject to

$$\sum_{i=1}^{N} v_i = 0 \tag{16}$$

$$\sum_{i=1}^{N} X_i v_i = 0 \tag{17}$$

$$-w_i \le v_i \le w_i, \quad i = 1, 2, \ldots, N . \tag{18}$$

Alternatively, by means of the substitution $t_i = v_i + w_i$ the dual problem can be written as

$$\text{maximize} \quad \sum_{i=1}^{N} Y_i t_i - \sum_{i=1}^{N} Y_i w_i \tag{19}$$

subject to

$$\sum_{i=1}^{N} t_i = \sum_{i=1}^{N} w_i \tag{20}$$

$$\sum_{i=1}^{N} X_i t_i = \sum_{i=1}^{N} X_i w_i \tag{21}$$

$$0 \le t_i \le 2w_i, \quad i = 1, 2, \ldots, N . \tag{22}$$

As before the dual problems (15)-(18) and (19)-(22) are bounded variable problems. In the case of problem (19)-(22) the variables are non-negative and upper bounded. It will be appropriate to examine the nature of optimal solutions to these problems as provided by the theory applicable to them (see Simonnard [10], Chapter 10 or Taha [11], Chapter 6).

First to be considered will be what is called an extremal solution. An extremal solution in the case of problem (15)-(18) is one in which two variables, say v_j and v_k, are selected to be basic and the other $N - 2$ variables are to be non-basic and set equal to one or the other of their extremal feasible values. Associated with the basic variables are the basic matrix $B = \begin{bmatrix} 1 & 1 \\ X_j & X_k \end{bmatrix}$ and the vectors $Y_B = [Y_j, Y_k]$ and v_B = column $[v_j, v_k]$. Index sets I_U and I_L may be defined such that $v_i = w_i$ if $i \in I_U$ and $v_i = -w_i$ if $i \in I_L$. Equations (16) and (17) then lead to the equations

$$v_j + v_k = \Sigma_{i \in I_L} w_i - \Sigma_{i \in I_U} w_i \tag{23}$$

$$X_j v_j + X_k v_k = \Sigma_{i \in I_L} X_i w_i - \Sigma_{i \in I_U} X_i w_i . \tag{24}$$

It will be convenient to let $A_U = \Sigma_{i \in I_U} w_i$, $A_L = \Sigma_{i \in I_L} w_i$, $C_U = \Sigma_{i \in I_U} X_i w_i$ and $C_L = \Sigma_{i \in I_L} X_i w_i$. Solving for v_B then produces

$$v_B = B^{-1} \begin{bmatrix} A_L - A_U \\ C_L - C_U \end{bmatrix} . \tag{25}$$

The extremal solution thus generated satisfies constraints (16) and (17). If the values for v_j and v_k given by equation (25) also satisfy constraint (18), then the solution is feasible. If a feasible extremal solution also satisfies the optimality conditions, then it is an optimal solution.

The optimality conditions for problem (15)-(18) are

$$Y_B B^{-1} \begin{bmatrix} 1 \\ X_i \end{bmatrix} - Y_i \leq 0 \quad \text{for} \quad i \in I_U , \tag{26}$$

and

$$Y_B B^{-1} \begin{bmatrix} 1 \\ X_i \end{bmatrix} - Y_i \geq 0 \quad \text{for} \quad i \in I_L . \tag{27}$$

Since $B = \begin{bmatrix} 1 & 1 \\ X_j & X_k \end{bmatrix}$, $B^{-1} = \frac{1}{X_k - X_j} \begin{bmatrix} X_k & -1 \\ -X_j & 1 \end{bmatrix}$ and

$$Y_B \cdot B^{-1} \begin{bmatrix} 1 \\ X_i \end{bmatrix} = Y_j \frac{X_k - X_i}{X_k - X_j} + Y_k \frac{X_i - X_j}{X_k - X_j} . \tag{28}$$

The latter quantity is the ordinate corresponding to X_i on the line in the (X, Y) plane which passes through the points (X_j, Y_j) and (X_k, Y_k). Condition (26) is that (X_i, Y_i) lies above that line and condition (27) is that (X_i, Y_i) lies below that line. A method of generating

extremal solutions which satisfy the optimality conditions of problem (15)-(18) is thus apparent: select a pair of points, (X_j, Y_j) and (X_k, Y_k) from the given set in the (X, Y) plane and let v_j and v_k be the basic variables; for the points (X_i, Y_i) which lie above the line set the corresponding dual variables $v_i = w_i$, for the remaining points lying on or below the line set the corresponding dual variables $v_i = -w_i$. The values of v_j and v_k are then determined from equation (25).

Of course the dual problem optimality conditions are nothing other than the feasibility conditions of the primal. An extremal dual solution which satisfies the optimality conditions will be called primal feasible. An extremal dual solution which is primal feasible and dual feasible, i.e. satisfies constraint (18), is an optimal solution.

4. Sharpe's Algorithm

Sharpe proceeds by parameterizing on b. With b held constant at $b = b'$ the problem is then to minimize $S(a, b') = \sum_{i=1}^{N} w_i |Y_i - (a + b'X_i))$. Let $Y_i' = Y_i - bX_i'$ so that

$$S(a, b') = \sum_{i=1}^{N} w_i |Y_i' - a| \; .$$

With $P_i - N_i = Y_i' - a$ the primal linear programming problem is

$$\text{minimize} \quad \sum_{i=1}^{N} w_i(P_i + N_i)$$

subject to

$$P_i - N_i + a = Y_i'$$

$$P_i \geq 0, \; N_i \geq 0, \qquad i = 1, 2, \ldots, N \; .$$

The dual problem is

$$\text{maximize} \quad \sum_{i=1}^{N} Y_i' v_i \tag{29}$$

subject to

$$\sum_{i=1}^{N} v_i = 0 \tag{30}$$

$$-w_i \leq v_i \leq w_i, \qquad i = 1, 2, \ldots, N \; , \tag{31}$$

or alternatively with $t_i = v_i + w_i$

$$\text{maximize} \quad \sum_{i=1}^{N} Y_i' t_i - \sum Y_i' w_i \tag{32}$$

subject to

$$\sum_{i=1}^{N} t_i = \sum_{i=1}^{N} w_i \tag{33}$$

$$0 \leq t_i \leq 2w_i, \qquad i = 1, 2, \ldots, N \; . \tag{34}$$

As Rao and Srinivasan [5] have noted the latter is a knapsack problem without integer restriction and has as its solution $a = Y'_j$ where Y'_j is the median among the values of Y'_i relative to the weights w_i. From the point of view of bounded variable linear programming theory that means starting with the largest value of Y'_i set the corresponding dual variable $t_i = 2w_i$ and assign i to the index set I_U. Continue similarly with successively smaller values of Y'_i as long as $2A_U \leq \sum_{i=1}^{N} w_i = W$. An index j will eventually be reached such that $2(A_U + w_j) > W$. Then set $t_j = W - 2A_U$. Clearly the constraint $0 \leq t_j \leq 2w_j$ is then satisfied. For any remaining index i set $t_i = 0$ and assign i to I_L. With $A_U + w_j + A_L = W$, it follows that

$$t_j = A_L - A_U + w_j \tag{35}$$

and that the constraint $0 \leq t_j \leq 2w_j$ is equivalent to the two inequalities

$$A_U \leq A_L + w_j \tag{36}$$

$$A_U \geq A_L - w_j \tag{37}$$

from which follows the interpretation of Y'_j as the median among the values of Y'_i relative to the weights w_i.

Thus, at optimum problem (32)-(34) has t_j as the single basic variable. For indices assigned to I_U the corresponding dual variables are non-basic and set equal to their upper bounds, and for indices assigned to I_L the corresponding dual variables are non-basic and set equal to their lower bounds. If Y'_j is unique among the values of Y'_i, i.e. $Y'_j \neq Y'_\ell$ for any other index ℓ, then there is an interval of values of b about b' such that t_j continues as the basic variable and the optimal value of a lies on the basic line $a = Y_j - bX_j$. Sharpe calls that line the border for those values of b. Let $S_2(b) = \min_a \{S(a,b)\}$. Then over that interval $Y_i - bX_i > Y_j - bX_j$ for $i \in I_U$ and

$Y_i - bX_i < Y_j - bX_j$ for $i \in I_L$.

Hence,

$$
\begin{aligned}
S_2(b) &= S(Y_j - bX_j, b) \\
&= \sum_{i=1}^{N} w_i \left| Y_i - Y_j - b(X_i - X_j) \right| \\
&= \Sigma_{i \in I_U} w_i [Y_i - Y_j - b(X_i - X_j)] - \Sigma_{i \in I_L} w_i [Y_i - Y_j - b(X_i - X_j)] \\
&= D_U - D_L - b(C_U - C_L) - (Y_j - bX_j)(A_U - A_L)
\end{aligned}
\tag{38}
$$

where

$$D_U = \Sigma_{i \in I_U} w_i Y_i \quad \text{and} \quad D_L = \Sigma_{i \in I_L} w_i Y_i \,.$$

Thus, over the interval of values for which t_j remains basic $S_2(b)$ is linear in b with slope

$$S_2'(b) = C_L - C_U - X_j(A_L - A_U) \,. \tag{39}$$

It may be noted that formula (39) also follows from the interpretation of the dual variables at optimum as the rate of change of the objective function with respect to the right hand side quantities of the associated constraints.

Over its entire domain $S_2(b)$ is piecewise linear and convex with changes in slope occurring whenever the border intersects another basic line. For a value of b corresponding to such an intersection $S_2(b)$ is not differentiable and formula (39) does not hold. Sharpe points out that if for $b = b'$ the slope of $S_2(b)$ is negative, the optimal value of b is greater than b'. Vice versa if the slope is positive. The algorithm consists in following the border to an intersection with another basic line $a = Y_i - bX_i$ at which the slope of $S_2(b)$ changes sign. The coordinates of that point of intersection are the optimal values of a and b. Each time the border reaches another basic line it must be determined whether the border follows the new basic line or continues along the old line. That problem will now be considered.

Suppose b' corresponds to the point of intersection of the two basic lines $a = Y_j - bX_j$ and $a = Y_k - bX_k$, i.e. $Y_j - b'X_j = Y_k - b'X_k$. Consider an extremal solution to problem (19)-(22) with v_j and v_k as basic variables and

$$A_U + A_L + w_j + w_k = W . \qquad (40)$$

Two cases must be considered. First suppose $X_k > X_j$. Then for $b < b'$, (X_k, Y_k) lies above the line through (X_j, Y_j) with slope b and, hence, $k \in I_U$. In terms of the values at $b = b'$ conditions (36) and (37) for $a = Y_j - bX_j$ to be the border are

$$A_U + w_k \le A_L + w_j \qquad (41)$$

$$A_U + w_k \ge A_L - w_j, \quad \text{or} \qquad (42)$$

$$-(w_j + w_k) \le A_U - A_L \le w_j - w_k . \qquad (43)$$

For $b > b'$, index k transfers from I_U to I_L and the conditions become

$$A_U \le A_L + w_k + w_j \quad \text{and}$$

$$A_U \ge A_L + w_k - w_j, \quad \text{or}$$

$$w_k - w_j \le A_U - A_L \le w_j + w_k . \qquad (44)$$

In order for $a = Y_j - bX_j$ to be the border on both sides of the intersection both (43) and (44) must hold, or

$$-(w_j - w_k) \le A_U - A_L \le w_j - w_k . \qquad (45)$$

Next suppose $X_k < X_j$. Then for $b < b'$, (X_k, Y_k) lies below the line through (X_j, Y_k) with slope b and hence $k \in I_L$. In terms of values at $b = b'$ the conditions for $a = Y_j - bX_j$ to be the border line are

$$A_U \le A_L + w_k + w_j \quad \text{and}$$

$$A_U \ge A_L + w_k - w_j \quad \text{or}$$

$$w_k - w_j \le A_U - A_L \le w_k - w_j . \qquad (46)$$

For $b > b'$ the condition is

$$-(w_j + w_k) \leq A_U - A_L \leq w_j - w_k \, . \tag{47}$$

One additional result is required in order to implement Sharpe's algorithm namely Formula (39) expressed in terms of values at $b = b'$, i.e. when equation (40) holds. For $b < b'$ and $x_k > X_j$, (X_k, Y_k) lies above the line through (X_j, Y_j) with slope b. Hence along the border $a = Y_j - bX_j$,

$$S_2'(b) = C_L - C_U - X_j(A_L - A_U) - w_k(X_k - X_j) \, .$$

Formulas for the other combinations of conditions may be obtained in a similar manner.

The following table summarizes the results required to implement Sharpe's algorithm at the point of intersection of the two basic lines $a = Y_j - bX_j$ and $a = Y_k - bX$ in the (a, b) plane.

Conditions	Border Line	$S_2'(b)$
$b < b'$, $X_k > X_j$, $-(w_j+w_k) \leq A_U - A_L \leq w_j - w_k$	$a = Y_j - bX_j$	$C_L - C_U - X_j(A_L - A_U) - w_k(X_k - X_j)$
$w_j - w_k \leq A_U - A_L \leq w_j + w_k$	$a = Y_k - bX_k$	$C_L - C_U - X_k(A_L - A_U) - w_j(X_k - X_j)$
$b > b'$, $X_k > X_j$, $-(w_j+w_k) \leq A_U - A_L \leq w_k - w_j$	$a = Y_k - bX_k$	$C_L - C_U - X_k(A_L - A_U) - w_j(X_j - X_k)$
$w_k - w_j \leq A_U - A_L \leq w_j + w_k$	$a = Y_j - bX_j$	$C_L - C_U - X_j(A_L - A_U) - w_k(X_j - X_k)$

5. Karst's Algorithm

Karst proceeds by selecting an arbitrary line in the (a, b) plane and following it until the minimum of $S(a, b)$ above it is found. The initially selected line may or may not be a basic line, but, it will be seen, that minimum does occur above the intersection of the initially selected line with one of the basic lines. The algorithm continues by following the just encountered basic line until the minimum above it is found, again at the intersection with a basic line. The process continues until the basic line just found is not a new one but one previously encountered. The last encountered point of intersection of basic lines has coordinates which are optimal values of a and b.

In terms of the (X, Y) plane the algorithm is equivalent to selecting an arbitrary initial point and solving the restricted problem of minimizing $S(a, b)$ over the set of lines which pass through that point. An efficient procedure for solving the restricted problem is the key element in Karst's algorithm, because each iteration requires the solution of just such a problem.

One linear programming formulation of the restricted problem may be obtained as a special case of problems (15)-(18) or (19)-(22) merely by the elimination of one constraint. The argument is as follows. Suppose that the initially selected point is one of the original data points (X_j, Y_j). (If it is not, include it as the $(N + 1)$st point with $w_{N+1} = 0$.) Then the restricted problem is obtained as a special case merely by making w_j very large in the primal problem, because then an optimal solution must pass through (X_j, Y_j). In the dual problems when w_j is very large the constraints $-w_j \leq v_j \leq w_j$ and $0 \leq t_j \leq 2w_j$ become non-restrictive and thus may be eliminated.

Hence, an extremal solution of problem (15)-(18) which has v_j and v_k basic and which is primal feasible is optimal for the restricted problem if

$$-w_k \leq v_k \leq w_k \, . \tag{48}$$

However, from formula (25)

$$v_k = \frac{C_L - C_U - X_j(A_L - A_U)}{X_k - X_j} . \tag{49}$$

If $X_k > X_j$, then (48) is equivalent to

$$-w_k(X_k - X_j) \leq C_L - C_U - X_j(A_L - A_U) \leq w_k(X_k - X_j), \quad \text{or} \tag{50}$$

$$C_U - X_jA_U - w_k(X_k - X_j) \leq C_L - X_jA_L \leq C_U - X_jA_U + w_k(X_k - X_j) . \tag{51}$$

With $x_i = X_i - X_j$ and $c_U = \Sigma\, i \in I_U w_i x_i$ and $c_L = \Sigma\, i \in I_L w_i x_i$, formula (49) becomes

$$v_k = \frac{c_L - c_U}{x_k} ,$$

and inequalities (51) may be written

$$c_U - w_k x_k \leq c_L \leq c_U + w_k x_k . \tag{52}$$

If $X_k < X_j$, then (48) is equivalent to

$$c_U + w_k x_k \leq c_L \leq c_U - w_k x_k . \tag{53}$$

Thus, a solution to the Karst restricted problem may be obtained by inspecting the extremal solutions to problem (15)-(18) which are primal feasible until one is found which satisfies either condition (52) or condition (53). The difficulty is that of generating those solutions in a convenient and systematic way. One approach to overcoming that difficulty will now be developed. Given that $a = Y_j - bX_j$

$$\begin{aligned} Y_i - (a + bX_i) &= Y_i - (Y_j - bX_j + bX_i) \\ &= y_i - bx_i . \end{aligned}$$

The conditions under which $y_i > bx_i$ and index i therefore assigned to I_U are either $b < \frac{y_i}{x_i}$ and $x_i > 0$ or $b > \frac{y_i}{x_i}$ and $x_i < 0$. Similarly the conditions under which $y_i < bx_i$ and index i is assigned to I_L are either $b < \frac{y_i}{x_i}$ and $x_i < 0$ or $b > \frac{y_i}{x_i}$ and $x_i > 0$.

The various situations may be visualized in terms of a plot of the given points in the (x, y) plane and a line with slope b passing through the origin in that plane.

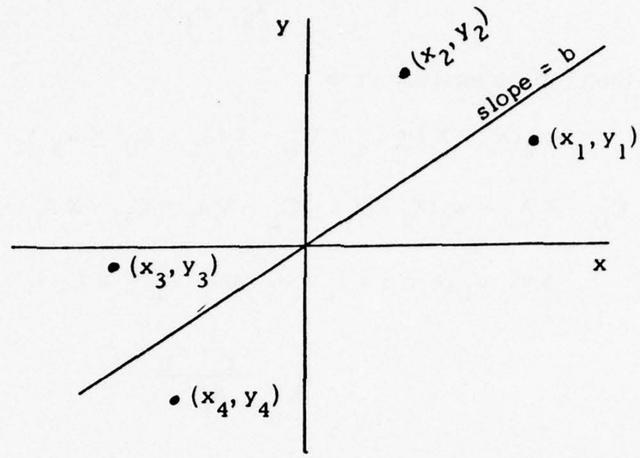

Figure 2

Index 2 is assigned to I_U because $b < \frac{y_2}{x_2}$ and $x_2 > 0$; index 3 is assigned to I_U because $b > \frac{y_3}{x_3}$ and $x_3 < 0$. On the other hand index 1 is assigned to I_L because $b > \frac{y_1}{x_1}$ and $x_1 > 0$ and index 4 is assigned to I_L because $b < \frac{y_4}{x_4}$ and $x_4 < 0$. The foregoing leads to the following procedure for solving the Karst restricted problem:

1) Compute x_i and y_i for $i = 1, 2, \ldots, N$, $i \neq j$. Assign i to I_U if $x_i > 0$ and i to I_L if $x_i < 0$. (If the X_i values are not all distinct, so that $x_i = 0$ is possible, assign i to I_U when $x_i = 0$ and $y_i > 0$ and i to I_L when $x_i = 0$ and $y_i < 0$).

2) Compute $\frac{y_i}{x_i}$ except when $x_i = 0$ and rank in ascending order.

3) For the smallest value of $\frac{y_i}{x_i}$ say $\frac{y_k}{x_k}$, if $x_k > 0$, delete k from I_U and test for satisfaction of condition (52). If (52) is satisfied, then $b = \frac{y_k}{x_k}$ is optimal. If

not, transfer k to I_L and proceed to the next higher value of $\frac{y_i}{x_i}$. If $x_k < 0$, delete k from I_L and test for satisfaction of condition (53). If (53) is satisfied, then $b = \frac{y_k}{x_k}$ is optimal. If not transfer k to I_U and proceed to next higher value of $\frac{y_i}{x_i}$.

4) Continue the process until an index k is found for which condition (52) or (53) is satisfied.

The solution to the restricted problem may be a solution to the unrestricted problem. From equation (25)

$$v_j = \frac{X_k(A_L - A_U) - (C_L - C_U)}{X_k - X_j} = \frac{x_k(A_L - A_U) - (C_L - C_U)}{x_k} . \tag{54}$$

If the value of v_j satisfies $-w_j \leq v_j \leq w_j$, then an optimal solution to the overall problem has been found.

Karst approaches the problem from a point of view not directly based upon a linear programming formulation. He noted that

$$G(b) = S(Y_j - bX_j, b) = \sum_{i=1}^{N} w_i |y_i - bx_i|$$

is piecewise linear and convex as a function of b. Furthermore the individual term $w_i |y_i - bx_i|$ has slope $-w_i |x_i|$ for $b < \frac{y_i}{x_i}$ and slope $w_i |x_i|$ for $b > \frac{y_i}{x_i}$. Hence $G(b)$ has slope $-\sum_{i=1}^{N} w_i |x_i|$ for $b < \min_i \left\{\frac{y_i}{x_i}\right\}$ and slope $\sum_{i=1}^{N} w_i |x_i|$ for $b > \max_i \left\{\frac{y_i}{x_i}\right\}$, and the slope of $G(b)$ increases by $2w_i |x_i|$ as b increases from $\frac{y_i}{x_i} - \epsilon$ to $\frac{y_i}{x_i} + \epsilon$ where $\epsilon > 0$. The index k is sought for which the slope of $G(b)$ changes from negative to positive. Thus by ranking the values of $\frac{y_i}{x_i}$ in ascending order and successively adding $2w_i |x_i|$ to an initial value of $-\sum_{i=1}^{N} w_i |x_i|$ the index k may be found which changes the total to a positive quantity. As before the solution is then $b = \frac{y_k}{x_k}$. Karst proceeds

with successive applications of the procedure until the index k so obtained is one previously encountered. He does not take advantage of the stopping rule of the simplex or dual simplex algorithm which is what equation (54) and the following inequality accomplish. It is suggested that the recognition of that stopping rule represents an improvement to the Karst algorithm.

An alternative linear programming formulation of the restricted problem can be developed. If $a = Y_j - bX_j$ is substituted in equation (13), the primal problem that results is

$$\text{minimize} \quad \sum_{i=1}^{N} w_i(P_i + N_i)$$

subject to

$$P_i - N_i + bx_i = y_i$$

$$P_i \geq 0, \; N_i \geq 0, \quad i = 1, 2, \ldots, N \, .$$

The dual problem is

$$\textit{maximize} \quad \sum_{i=1}^{N} y_i v_i$$

subject to

$$\sum_{i=1}^{N} x_i v_i = 0$$

$$-w_i \leq v_i \leq w_i, \quad \text{or}$$

with $t_i = v_i + w_i$

$$\text{maximize} \quad \sum_{i=1}^{N} y_i t_i - \sum_{i=1}^{N} y_i w_i$$

subject to

$$\sum_{i=1}^{N} x_i t_i = \sum_{i=1}^{N} x_i w_i, \quad 0 \leq t_i \leq 2w_i, \quad i = 1, 2, \ldots, N \, .$$

The latter problem appears to be a knapsack problem for which a median type solution is available. However, when some of the x_i values are positive and some negative, the simple median procedure does not apply. Thus, the alternative linear programming formulation does not appear to be any more useful than the previous formulation.

Barrodale and Roberts [3] have proposed an algorithm for the general discrete ℓ_1 linear approximation problem. Their algorithm is based on the primal formulation (minimization) and takes advantage of the special structure of that formulation. They observe that many columns in the full simplex tableau need not be explicitly retained and may be inferred from the critical columns that are retained. If the values of n parameters are to be determined (two in the case of fitting a line), then only n columns need to be maintained from one iteration to another. Furthermore they have devised a way of combining several simplex iterations into one iteration of their algorithm. It will be seen in the examples which follow that the Karst algorithm when $(0,0)$ is the initial point in the (X,Y) plane is identical to the Barrodale and Roberts procedure in terms of path in the (a,b) plane followed to optimum.

6. Example

Example 1: A five point problem with values given in the table at the right.

i	X_i	Y_i	w_i
1	1	1	1
2	2	5	2
3	3	10	6
4	4	22	3
5	5	18	2
			$W = 14$

Sharpe's Algorithm

First minimization along $b = 0$.

i	Y'_i	Rank	$\dfrac{Y_i - Y_3}{X_i - X_3}$
1	1	5	4.5
2	5	4	5.0
3	10	3	—
4	22	1	12.0
5	18	2	4.0

$W = 14$

$-2w_4 = \underline{-6}$
$\ 8$

$-2w_5 = \underline{-4}$
$\ 4$

$-2w_3 = \underline{-12}$
$\ -8$

$j = 3$, $a = 10$ and $a = 10 - 3b$ is the border line for an interval of values b near $b = 0$. $I_U = \{4, 5\}$, $I_L = \{1, 2\}$, $A_U = 5$, $A_L = 3$, $C_U = 22$, $C_L = 5$, $D_U = 102$ and $D_L = 1$.

Hence $S_2(b) = D_U - D_L - Y_3(A_U - A_L) - b[C_U - C_L - X_3(A_U - A_L)] = 71 - 11b$. Since $S'_2(b) = -11$, b must be increased. The value of b for which the border line first intersects with another basic line is $b = 4$ at the intersection with $a = 18 - 5b$. At this point $j = 3$, $k = 5$, $I_U = \{4\}$, $I_L = \{1, 2\}$, $A_U = 3$, $A_L = 3$, $C_U = 12$, $C_L = 5$, $D_U = 66$ and $D_L = 11$. With $w_j = 6$ and $w_k = 2$, the condition which holds is $w_k - w_j \leq A_U - A_L \leq w_k + w_j$. Hence, $a = 10 - 3b$ continues as the border for $b > 4$. Also, for $b > 4$, index $5 \in I_L$ and $S_2(b) = 39 - 3b$. Thus $b = 4$ and $a = -2$ is not optimal.

The next critical value of b is b = 4.5 which occurs for k = 1 at the intersection with $a = 1 - b$. $I_U = \{4\}$, $I_L = \{2,5\}$, $A_U = 3$, $A_L = 4$, $C_U = 12$, $C_L = 14$, $D_U = 66$ and $D_L = 46$. Since $w_k - w_j \leq A_U - A_L \leq w_k + w_j$ holds, $a = 10 - 3b$ continues as the border for $b > 4.5$. However, for $b > 4.5$, index $1 \in I_U$ and

$$S_2(b) = 21 + b\,,$$

which means the optimal solution has been found: $b = 4.5$, $a = -3.5$ and $S(a,b) = 25.5$.

<u>Karst's Algorithm</u>

Let the initial fixed point in the (X, Y) plane be the origin.

Iteration 1: Minimize along $a = 0$.

i	x_i	y_i	w_i	$\frac{y_i}{x_i}$
1	1	1	1	1
2	2	5	2	2.5
3	3	10	6	3.3
4	4	22	3	4.5
5	5	18	2	3.6

i	C_U	C_L	Condition (52) or (53) satisfied?
	45	0	
	-1	—	
1	44	0	no
	—	1	
	44	1	
	-4	—	
2	40	1	no
	—	4	
	40	5	
	-18	—	
3	22	5	yes

k = 3

Iteration 2: Along $a = 10 - 3b$ with $j = 3$.

i	x_i	y_i	$\frac{y_i}{x_i}$
1	-2	-9	4.5
2	-1	-5	5
3	0	0	-
4	1	12	12
5	2	8	4

i	$w_i x_i$	C_U	C_L	Conditions (52) or (53) satisfied?
		7	-4	
5	4	-4	—	
		3	-4	no
		—	+4	
		3	0	
1	-2	—	-(-2)	
		3	2	yes

$$\therefore\ k = 1$$

$$v_1 = \frac{C_L - C_U}{x_1} = \frac{2 - 3}{-2} = \frac{1}{2}$$

$$v_3 = \frac{x_1(A_L - A_U) - (C_L - C_U)}{x_1} = \frac{1}{2}$$

Since $-w_1 \leq v_1 \leq w_1$, and $-w_3 \leq v_3 \leq w_3$, the solution is optimal with $b = 4.5$ and $a = -3.5$ as before.

Barrodale and Roberts Improved Algorithm

The full primal linear programming version of this problem requires fourteen variables and five constraints after the variables unrestricted in sign a and b are each replaced by the difference of two non-negative variables. However, as Barrodale and Roberts have shown the normal simplex tableaux contain many columns which need not be maintained. Only two columns corresponding to non-basic variables plus the column of current solution values are needed. Again starting with $(a, b) = (0, 0)$ the sequence of condensed tableaux is shown below.

Initial Tableau

Basis	Costs	a	b	R
P_1	1	1	1^*	1
P_2	2	1	2	5
P_3	6	1	3	10
P_4	3	1	4	22
P_5	2	1	5	18
Marginal Costs		14	45	173

The marginal cost for b is greatest so that b will become basic at the first iteration. The normal simplex pivot (1^*) corresponds to $b = 1$, but b can be increased beyond that value by making P_1 non-basic and N_1 basic. That is accomplished by subtracting $2w_1$ from the marginal cost row. In a similar manner P_2 will be replaced by N_2 as b increases beyond 2.5. Finally b will become basic replacing P_3 bringing the first iteration to an end. The marginal cost row went through the following sequence of values as N_1 and N_2 became basic.

	a	b	R
Initial tableau	14	15	173
N_1 replaces P_1	12	43	171
N_2 replaces P_2	8	35	151

Notice that if the value of b goes beyond $\frac{10}{3}$ so that N_3 replaces P_3, the marginal cost for b becomes negative which means that b should not increase beyond $\frac{10}{3}$ at this iteration.

Second Tableau

Basis	Costs	-a	P_3	R	Critical Ratios
N_1	1	$+\frac{2}{3}$**	$\frac{1}{3}$	$\frac{7}{3}$	3.5
N_2	2	$+\frac{1}{3}$	$\frac{2}{3}$	$\frac{5}{3}$	5
b	0	$-\frac{1}{3}$	$\frac{1}{3}$	$\frac{10}{3}$	-10
P_4	3	$+\frac{1}{3}$	$-\frac{4}{3}$	$\frac{26}{3}$	26
P_5	2	$+\frac{2}{3}$*	$-\frac{5}{3}$	$\frac{4}{3}$	2
Marginal Costs		$\frac{11}{3}$	$-\frac{35}{3}$	$\frac{103}{3}$	

The second tableau does not represent an optimal solution because the marginal cost of $-a$ is positive indicating that decreasing a below 0 will decrease total cost. The normal simplex pivot $(\frac{2}{3}*)$ corresponds to $-a = 2$, but a can be decreased further by making N_5 basic replacing P_5. The marginal cost row then becomes

	$-a$	P_3	R
Marginal costs	1	-5	29

The pivot then employed $(\frac{2}{3}**)$ brings the tableau to the following:

Third Tableau

Basis	Costs	P_1	P_3	R
$-a$	0	$\frac{3}{2}$	$\frac{1}{2}$	3.5
N_2	2	$-\frac{1}{2}$	$\frac{1}{2}$	.5
b	0	$\frac{1}{2}$	$\frac{1}{2}$	4.5
P_4	3	$-\frac{1}{2}$	$-\frac{3}{2}$	7.5
N_5	2	1	2	1
Marginal Costs		$-\frac{1}{2}$	$-\frac{11}{2}$	25.5

The marginal costs of P_1 and P_3 are negative, and the marginal costs of N_1 and N_3 can both be inferred to be $-\frac{1}{2}$ from the relation: sum of marginal costs of P_i and N_i equals $-2w_i$. Therefore, the solution represented by the third tableau is optimal with $a = -3.5$, $b = 4.5$ and $S(a, b) = 25.5$.

Comparison of the Barrodale and Roberts condensed tableau with the Karst algorithm reveals that in this example they are identical in terms of the sequence of values of a and b through which they progressed to the optimal solution.

Example 2: An eleven point problem with values as shown.

i	X_i	Y_i	w_i
1	1	34	3
2	2	24	5
3	3	31	8
4	4	40	10
5	5	30	15
6	6	49	20
7	7	48	23
8	8	48	20
9	9	67	15
10	10	58	13
11	11	67	11
			W = 143

Sharpe's Algorithm

				b = 0					
i	X_i	Y_i	w_i	Y'_i	Rank	U / L	$\frac{Y_i - Y_7}{X_i - X_7}$	$\frac{Y_i - Y_4}{X_i - X_4}$	$\frac{Y_i - Y_{10}}{X_i - X_{10}}$
1	1	34	3	34	8	L	2.33	2	2.67
2	2	24	5	24	11	L	4.80	8	4.25
3	3	31	8	31	9	L	4.25	9	3.86
4	4	40	10	40	7	L	2.67	-	3
5	5	30	15	30	10	L	9.00	-10	5.6
6	6	49	20	49	4	U	-1.00	4.5	2.25
7	7	48	23	49	5-6	-	—	2.67	3.33
8	8	48	20	48	5-6	-	0	2	5
9	9	67	15	67	1-2	U	9.50	5.4	-9
10	10	58	13	58	3	U	3.33	3	-
11	11	67	11	67	1-2	U	4.75	3.86	9
			143						

i	w_i	k	w_k	a	b'	A_U	A_L	C_U	C_L	Slope	Changes	S
7	23	8	20	48	0	59	41	506	152			1292
b should increase and for $b > 0$, 7 is the border index.										-208	$8 \in I_L$	
7	23	1	3	31.67	2.33	59	58	506	309		$1 \notin I_L$	806.67
b should increase and for $b > 2.33$, 7 is the border index.										-172	$1 \in I_U$	
7	23	4	10	29.33	2.67	62	48	509	269		$4 \notin I_L$	749.33
b should increase and for $b > 2.67$, 4 is the border index.										-115	$7 \in I_L$	
4	10	10	13	28	3.0	49	71	379	430		$10 \notin I_U$	711
b should increase and for $b > 3.0$, 10 is the border index.										-109	$4 \in I_U$	
10	13	7	23	24.67	3.33	59	48	419	269		$7 \notin I_L$	674.67
b should increase and for $b > 3.37$, 7 is the border index.										- 34	$10 \notin I_L$	
7	23	3	8	18.25	4.25	59	53	419	375		$3 \notin I_L$	643.5
for $b > 4.25$, 3 is the border index, but b should not increase.										154		

Thus, the optimal solution is $a = 18.25$, $b = 4.25$ and $S(a,b) = 643.50$.

Karst's Algorithm

Iteration 1: Minimize along $a = 0$.

i	w_i	x_i	y_i	$\frac{y_i}{x_i}$	
1	3	1	34	34	
2	5	2	24	12	
3	8	3	31	10.33	
4	10	4	40	10	
5	15	5	30	6	3
6	20	6	49	8.17	
7	23	7	48	6.86	5
8	20	8	48	6	2
9	15	9	67	7.44	
10	13	10	58	5.8	1
11	11	11	67	6.09	4
12	0	0	0	—	

$-\Sigma\, w_i |x_i| = -979$

$2w_{10}|x_{10}| = \underline{260}$
-719

$2w_8|x_8| = \underline{320}$
-399

$2w_5|x_5| = \underline{150}$
-249

$2w_{11}|x_{11}| = \underline{242}$
$-\ 7$

$2w_7|x_7| = \underline{322}$
$+315$

$\therefore\ k = 7$

Iteration 2: Along $a = 48 - 7b$, $j = 7$.

x_i	y_i	$\frac{y_i}{x_i}$	U/L
-6	-14	2.33	U
-5	-24	4.80	L
-4	-17	4.25	
-3	- 8	2.67	U
-2	-18	9.00	L
-1	1	-1.00	U
0	0	—	
1	0	0	L
2	19	9.50	U
3	10	3.33	L
4	19	4.75	U

$-\Sigma\, w_i |x_i| = -288$

$2w_6|x_6| = \underline{40}$
-248

$2w_8|x_8| = \underline{40}$
-208

$2w_1|x_1| = \underline{36}$
-172

$2w_4|x_4| = \underline{60}$
-112

$2w_{10}|x_{10}| = \underline{78}$
$-\ 34$

$2w_3|x_3| = \underline{64}$
30

$\therefore\ k = 3$

$A_L = 53,\ A_U = 59$

$C_L = 4,\ C_U = 6$

$v_3 = \frac{4 - 6}{-4} = .5$

$v_7 = A_L - A_U - v_3 = -6.5$

Optimal solution: $b = 4.25$

$a = 48 - 7(4.25) = 18.25$

The Barrodale and Roberts method will again follow the same path to the optimal solution as the Karst algorithm, and will not be shown here.

7. Conclusions

The Karst algorithm is similar to the Barrodale and Roberts method in that a single iteration of each is equivalent to several standard simplex iterations. The two will follow the same path in the (a, b) plane to optimum starting from $(0, 0)$ if initially a unit change in b produces a greater decrease in $S(a, b)$ than a unit change in a.

For the single purpose of finding the optimal values of a and b the Karst algorithm is very efficient and undoubtedly involves less computation than Sharpe's algorithm.

The Sharpe algorithm generates the optimal values of a for values of b near the optimal value of b, and from them the values of $S_2(b)$. If the function $S_2(b)$ is of interest, then the Sharpe algorithm may be preferred. Otherwise, it is more cumbersome than Karst's algorithm.

References

1. R. D. Armstrong and E. L. Frome: A Comparison of Two Algorithms for Absolute Deviation Curve Fitting. Journal of the American Statistical Association, 71 (June 1976), 328-30.
2. I. Barrodale and F. D. K. Roberts: An Improved Algorithm for Discrete ℓ_1 Linear Approximation, TSR 1172, Mathematics Research Center, Madison, Wis., 1972.
3. I. Barrodale and F. D. K. Roberts: An Improved Algorithm for Discrete ℓ_1 Linear Approximation. SIAM Journal of Numerical Analysis, 10 (October 1973), 839-48.
4. Otto J. Karst: Linear Curve Fitting Using Least Deviations. Journal of the American Statistical Association (March 1958), pp. 118-132.
5. M. R. Rao and V. Srinivasan: A Note on Sharpe's Algorithm for Minimizing the Sum of Absolute Deviations in a Simple Regression Problem. Management Science, Vol. 19, No. 2 (October, 1972), 222-25.
6. John R. Rice: The Approximation of Functions, Vol. I - Linear Theory. Addison-Wesley Publishing Company, Inc., Reading, Mass., 1964.
7. P. D. Roberts and A. Ben-Israel: An Interval Programming Algorithm for Discrete Linear L_1 Approximation Problems. Journal of Approximation Theory, 2 (December 1969), 323-36.
8. E. J. Schlossmacher: An Iterative Technique for Absolute Deviations Curve Fitting. Journal of the American Statistical Association, 68 (December 1973), 857-9.
9. W. F. Sharpe: Mean-Absolute Deviation Characteristic Lines for Securities and Portfolios. Management Science, Vol. 18, No. 2 (October 1971). pp. B1-B13.
10. Michel Simonnard: Linear Programming. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1966.

11. H. A. Taha: Operations Research, An Introduction. Second Edition, MacMillan Publishing Company, Inc., New York, N. Y., 1976.

12. H. M. Wagner: Linear Programming Techniques for Regression Analysis. Journal of the American Statistical Association, Vol. 50 (March 1959), pp. 206-212.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER (6) 1735	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER (9) Technical
4. TITLE (and Subtitle) A DEFENSE OF THE KARST ALGORITHM FOR FINDING THE LINE OF BEST FIT UNDER THE L_1 NORM.		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) Donald R. Schuette		8. CONTRACT OR GRANT NUMBER(s) (15) DAAG29-75-C-0024
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin, 610 Walnut Street, Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 5 (Mathematical Programming and Operations Research
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina 27709		12. REPORT DATE (11) March 1977
		13. NUMBER OF PAGES 31
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) (14) MRC-TSR-1735		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited. (12) 34p.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Linear curve fitting L_1 norm

Linear programming

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The methods proposed by Karst and Sharpe of fitting a line to a given set of points in the plane so as to minimize the sum of the absolute values of the deviations are examined by means of their linear programming formulations. It is seen that the Karst algorithm, which was criticized by Sharpe, already contains the type of improvement proposed by Barrodale and Roberts for ℓ_1 linear approximations. For the single purpose of finding the optimal line the Karst procedure appears to be very efficient. If, as Sharpe suggests may sometimes be the case, the investigator is interested in the sensitivity of the minimum sum to changes in the slope parameter, then Sharpe's algorithm is preferred. The Karst algorithm is improved by incorporating into it the simplex stopping rule. The problem is generalized to permit arbitrary weightings of the deviations.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

221 200